Redes Complexas

Uma introdução

Leandro Chaves Rêgo, Jesus Ossian Cunha, Pablo Ignacio Fierens

Universidade Federal do Ceará

05 e 06 de Novembro de 2024

Conteúdo

Introdução e Motivação

Redes complexas

Conceito

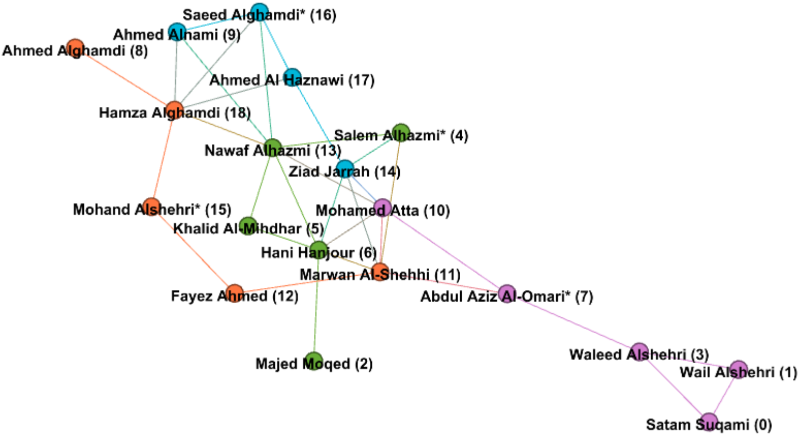
- São modelos utilizados para representar relações entre pares de elementos de um certo grupo de interesse.
- Se caracterizam pelo conjunto das relações não apresentarem um padrão de relações simples como o que ocorre no caso em que todos os pares de elementos se relacionam entre si, nem também se assemelham a relações que sejam formadas aleatoriamente.

Redes Complexas

Exemplos

- Internet.
- *World Wide Web*.
- Rede de comércio internacional.
- Rede de amizades no *Facebook*.
- Rede de seguidores no *Instagram*.
- Cadeia alimentar ecológica.
- Rede neural.

Rede dos terroristas no ataque de 11 de setembro de 2001



Fonte: Lindelauf et al. [2013]

Redes Complexas

Exemplos de aplicações

- Entender como notícias se espalham, como doenças se propagam.
- Procurar quem são os indivíduos mais importantes em uma rede de criminalidade.
- Identificar que produtos as pessoas tendem a comprar ou como elas tendem a votar.
- Analisar séries temporais com o objetivo de identificar padrões de anomalias em sinais biológicos ou verificação de assinaturas manuscritas online.

Propriedades Estruturais

Teoria de Grafos

Grafo

- Consiste de um par ordenado de conjuntos $G = (V, E)$.
- V é o conjunto de vértices ou nós do grafo, que representam os elementos que se relacionam em uma rede complexa.
- $E \subseteq V \times V$ é um conjunto de pares ordenados de vértices, chamados de arestas, que representam as relações entre os elementos.
- Se $(u, v) \in E$ implica que $(v, u) \in E$, então dizemos que o grafo é não direcionado ou não dirigido, caso contrário ele é chamado de direcionado, dirigido ou um digrafo.
- Uma aresta que liga um vértice de um grafo a ele mesmo é chamada de laço.
- Um subgrafo é qualquer grafo $G' = (V', E')$ tal que $V' \subseteq V$ e $E' \subseteq E$.

Teoria de Grafos

Matriz de adjacência

- A matriz de adjacência de um grafo em que $V = \{1, 2, \ldots, n\}$ tem seu elemento (i, j) dado por

$$A(i,j) = \begin{cases} 1, & \text{se } (i,j) \in E, \\ 0, & \text{caso contrário.} \end{cases}$$

- A matriz de adjacência de grafos não direcionados é simétrica.
- Se existir um laço ligando um nó i a ele mesmo, então $A(i,i) = 2$ se o grafo for não direcionado e igual a $A(i,i) = 1$ se o grafo for direcionado.

Teoria de Grafos

Grafos ponderados

- Algumas relações possuem uma intensidade maior que outras.
  - Em redes de comércio internacional: volume das transações.
  - Em relações de amizade: frequência de encontros.
- Utilizaremos uma matriz de adjacência ponderada, W, de modo que $A(i,j) = 0$ se e somente se $W(i,j) = 0$.
- Quanto maior $|W(i,j)|$, maior a intensidade da relação entre os nós.
- Pode ser uma relação positiva (por ex., amizade) ou negativa (por ex., inimizade).

Teoria de Grafos

Passeios

- Um passeio (em inglês, *walk*) consiste de uma sequência ordenada de vértices $(i_0, i_1, i_2, \ldots, i_k)$ tal que $(i_{j-1}, i_j) \in E$ para $j = 1, 2, \ldots, k$. Neste caso, k é o comprimento do passeio.
- O número de passeios de comprimento k entre i e j é dado por $A^k(i, j)$.
- No caso de redes ponderadas, como definimos que quanto maior o peso $W(i, j)$ mais intensa é a relação entre i e j, significa que eles são mais próximos. Deste modo, o comprimento de um passeio ou de um caminho é dado pela soma dos recíprocos dos pesos das arestas.

Teoria de Grafos

Caminhos

- Se um passeio não contém arestas repetidas, ele é conhecido como um caminho (em inglês, *path*).
- Um caminho sem vértices repetidos (exceto, possivelmente, pelo primeiro e último da sequência que podem ser iguais) é chamado de caminho simples.
- Um caminho de menor comprimento entre dois vértices i e j é chamado de caminho geodésico e o comprimento do caminho geodésico é conhecido como distância geodésica, denotada por $d(i, j)$.

Teoria de Grafos

Ciclos

- Se em um caminho $i_0 = i_k$, este caminho é conhecido como um ciclo.
- Se o caminho for simples e $i_0 = i_k$, temos um ciclo simples.
- Redes que não contém ciclos são chamadas de acíclicas.
  - Toda rede acíclica não contém laços.
  - Se a quantidade de vértices é finita, é possível rotular os vértices de uma rede direcionada acíclica de tal forma que a matriz de adjacência da rede é triangular com todos os elementos da diagonal principal iguais a zero.
  - Uma rede direcionada é acíclica se e somente se todos os autovalores da matriz de adjacência são nulos.

Teoria de Grafos

Conectividade

- Se existe um caminho entre dois vértices quaisquer da rede, dizemos que a rede é conectada.
- No caso de redes direcionadas:
  - Se a direção das arestas é levada em consideração ao observar o caminho, diz-se que a rede é fortemente conectada.
  - Caso contrário, temos que a rede é fracamente conectada.
- No caso em que a rede não seja conectada, alguns subgrafos podem ser conectados.
- Componente: é um subgrafo $G' = (V', E')$ conectado tal que não existe nenhum outro subgrafo do grafo original que seja conectado e cujos conjuntos de vértices e arestas contenham os conjuntos de vértices e arestas de G', respectivamente.

Teoria de Grafos

Vizinhança

- A vizinhança, $N_G(i)$, de um vértice i consiste de todos os vértices que são ligados a i.
- A k-ésima vizinhança, $N_G^k(i)$ de um nó i como o conjunto de nós que podem ser alcançados por caminhos de comprimento menor ou igual a k a partir de i.
- No caso de redes direcionadas:
  - Predecessores: $P_G(i) = \{j : A(j,i) = 1\}$.
  - Sucessores: $S_G(i) = \{j : A(i,j) = 1\}$.

Teoria de Grafos

Árvore e floresta

- Uma árvore é uma rede conectada e acíclica.
- Uma rede conectada com n vértices é uma árvore se e somente se ela possui $n-1$ arestas.
- Folhas: vértices que só possuem uma aresta.
- Linha: árvore que só possui duas folhas.
- Estrela: árvore que possui um vértice que toda aresta envolve este vértice.
- Uma floresta é uma rede em que cada componente é uma árvore.

Teoria de Grafos

Redes regulares

- Uma rede regular é uma rede em que cada vértice está envolvido no mesmo número de arestas.
- Círculo: cada vértice está envolvido em exatamente duas arestas.
- Rede completa: todas as arestas possíveis entre vértices distintos estão presentes.

Teoria de Grafos

Redes de Erdõs-Rényi

- Fixamos o número de vértices n e cada uma das $\binom{n}{2}$ possíveis arestas tem uma probabilidade $0 < p < 1$ fixa e independente de estar presente.
- Uma dada rede não ponderada que possua m arestas e n vértices tem probabilidade $p^m(1-p)^{\binom{n}{2}-m}$ de ser gerada.
- Apesar deste tipo de rede aleatória não se aproximar das características usualmente encontradas em redes complexas, é relativamente simples calcular métricas dessas redes e compará-las com redes complexas reais.

Grafos com Python

Python

Python é uma linguagem de programação interpretada, de alto nível e de propósito geral:

- *Interpretada*: há um interpretador, isto é, um programa que executa as instruções. O interpretador lê o código, o traduz a uma linguagem intermediária e o executa. Em contraposição, existem outras linguagens de programação em que os programas precisam ser compilados, i.e., traduzidos a código de máquina.
- *De alto nível*: Python tem um nível de abstração maior do que a linguagem de máquina, como é o caso da linguagem assembler.
- *De propósito geral*: Python não foi criado para uma aplicação em particular, como pode ser uma linguagem para o cálculo numérico ou para trabalhar com bases de dados.

Grafos com Python

Networkx

Um pacote de Python bem simples que permite trabalhar com redes
`https://networkx.org/`

Repositório do minicurso

`https://github.com/jesusossian/minicurso_sbpo2024`

Primeiro exemplo

Um exemplo clássico de rede social é um clube de karate constituído por 34 membros. As arestas do grafo representam amizades entre os membros. Este exemplo está incluso no pacote Networkx.

Clube de Karate I

```python
import networkx as nx                  #importa o pacote Networkx com o pseudônimo nx
g = nx.karate_club_graph()             #g é o grafo do clube de karate
print('====================')
print('Nós')
print(g.nodes())                       #mostra os nós: só números porque está anonimizado
print('====================')
print('Arestas')
print(g.edges())                       #mostra as arestas como pares de nós
```

```
#Resultados
====================
Nós
[0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33]
====================
Arestas
[(0, 1), (0, 2), (0, 3), (0, 4), (0, 5), (0, 6), (0, 7), (0, 8), (0, 10), (0, 11), (0, 12), (0, 13), (0, 17), (0, 19), (0, 21), (0, 31), (1, 2), (1, 3), (1, 7), (1, 13), (1, 17), (1, 19), (1, 21), (1, 30), (2, 3), (2, 7), (2, 8), (2, 9), (2, 13), (2, 27), (2, 28), (2, 32), (3, 7), (3, 12), (3, 13), (4, 6), (4, 10), (5, 6), (5, 10), (5, 16), (6, 16), (8, 30), (8, 32), (8, 33), (9, 33), (13, 33), (14, 32), (14, 33), (15, 32), (15, 33), (18, 32), (18, 33), (19, 33), (20, 32), (20, 33), (22, 32), (22, 33), (23, 25), (23, 27), (23, 29), (23, 32), (23, 33), (24, 25), (24, 27), (24, 31), (25, 31), (26, 29), (26, 33), (27, 33), (28, 31), (28, 33), (29, 32), (29, 33), (30, 32), (30, 33), (31, 32), (31, 33), (32, 33)]
```

Clube de Karate II

Vizinhos:

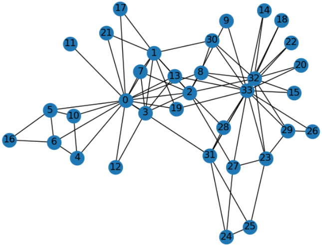
```
for n in g.neighbors(12):     #vizinhos do nó 12
    print(n)
```

```
#Resultados
0
3
```

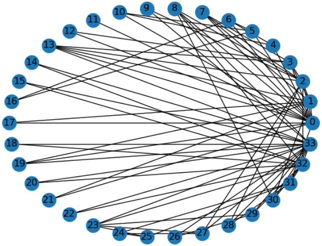
Gráficos:

```
nx.draw(g, with_labels = True)    #mostra graficamente a rede
nx.draw_circular(g, with_labels = True)
nx.draw_kamada_kawai(g, with_labels = True)
```

Clube de Karate III

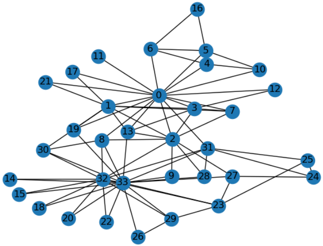


(a) Desenho aleatório



(b) Desenho circular

Clube de Karate IV



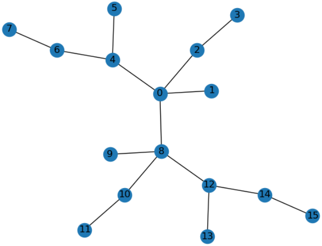
(c) Desenho de Kamada-Kawai

Algumas Redes Especiais I

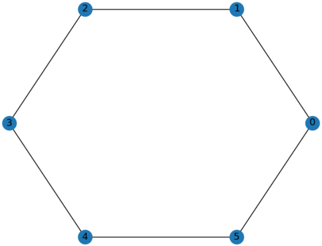
```
#árvore binomial
gab = nx.binomial_tree(4)
#círculo com 6 nós
gci = nx.circulant_graph(6,[1])
#grafo regular com 6 nós e 3 arestas por nó
gre = nx.random_regular_graph(3, 6)
#grafo completo com 6 nós
gco = nx.complete_graph(6)
#grafo aleatório de Erdös-Rényi com 6 nós e probabilidade 0.25
ger = nx.gnp_random_graph(6,0.25)
```

```
#Grafo direcionado
gfa = nx.DiGraph()
#Lista de nós
gfa.add_nodes_from(['Avó','Avô','Pai','Mãe','Filho','Filha'])
#Arestas
gfa.add_edge('Avó','Pai')     #Avó -> Pai
gfa.add_edge('Avô','Pai')     #Avô -> Pai
gfa.add_edge('Pai','Filho')   #Pai -> Filho
gfa.add_edge('Pai','Filha')   #Pai -> Filha
gfa.add_edge('Mãe','Filho')   #Mãe -> Filho
gfa.add_edge('Mãe','Filha')   #Mãe -> Filha
```

Algumas Redes Especiais II

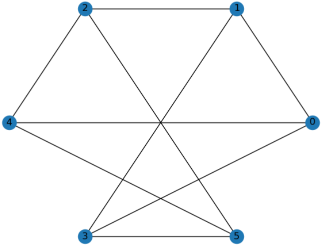


(a) Árvore

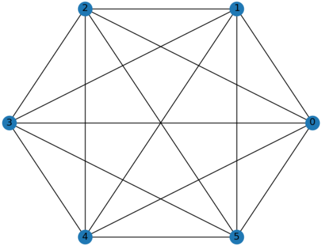


(b) Círculo com 6 nós

Algumas Redes Especiais III

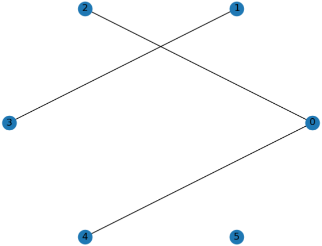


(c) Grafo regular com 6 nós e 3 arestas por nó

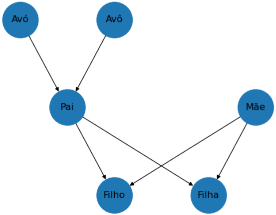


(d) Grafo completo com 6 nós

Algumas Redes Especiais IV



(e) Grafo aleatório de Erdös-Rényi com 6 nós e probabilidade 0.25



(f) Família

Métricas Globais

Densidade

Seja uma rede $G = (V, E)$, com $||V|| = n$ e m arestas. Então, a densidade da rede é:

$$\text{dens}(G) = \begin{cases} \frac{2m}{n(n-1)}, & \text{se a rede for não direcionada, e} \\ \frac{m}{n(n-1)}, & \text{se a rede for direcionada.} \end{cases}$$

```
g = nx.karate_club_graph()
#grafo regular com 6 nós e 3 arestas por nó
gre = nx.random_regular_graph(3, 6)
#grafo completo com 6 nós
gco = nx.complete_graph(6)
print(nx.density(g))
print(nx.density(gre))
print(nx.density(gco))
```

```
#Resultados
0.13903743315508021
0.6
1.0
```

Métricas Globais

Cumprimento de caminhos

- Diâmetro: maior distância geodésica entre dois nós em uma mesma componente.
- Comprimento médio dos caminhos: média de todas as distâncias geodésicas entre dois nós que são conectados na rede.

Métricas Globais

```
g = nx.karate_club_graph()
caminho = nx.shortest_path(g, 3, 20)
longitude = nx.shortest_path_length(g,3,20)
print(caminho)
print(longitude)
```

```
#Resultados
[3, 2, 32, 20]
3
```

```
#Comprimento médio dos caminhos
cm = nx.average_shortest_path_length(g)
#Diâmetro
di = nx.diameter(g)

print(cm)
print(di)
```

```
#Resultados
2.408199643493761
5
```

Métricas Globais

Grau

- Rede não direcionada:

$$d(i) = \sum_{j \in V} A(i, j).$$

- Rede direcionada:

$$d^{out}(i) = \sum_{j \in V} A(i, j)$$

$$d^{in}(i) = \sum_{j \in V} A(j, i).$$

Redes Ponderadas

- Em redes ponderadas, o grau ponderado é calculado substituindo a matriz de adjacência, A, pela matriz de adjacência ponderada, W.

Métricas Globais

Distribuição de graus

- A frequência relativa de ocorrência dos graus nos vértices da rede:

$$P(d) = \frac{||V_d||}{n},$$

em que $n = ||V||$, $V_d = \{i \in V : d(i) = d\}$ e $d = 0, 1, \ldots$

- No caso de uma rede binomial com probabilidade p de ocorrência de cada aresta:

$$P(d) = \binom{n-1}{d} p^d (1-p)^{n-1-d}.$$

- Em muitas redes complexas, é comum que a distribuição dos graus seja livre de escala, ou seja, siga uma distribuição de lei de potências:

$$P(d) = cd^{-\alpha}, \qquad c, \alpha > 0.$$

Métricas Globais

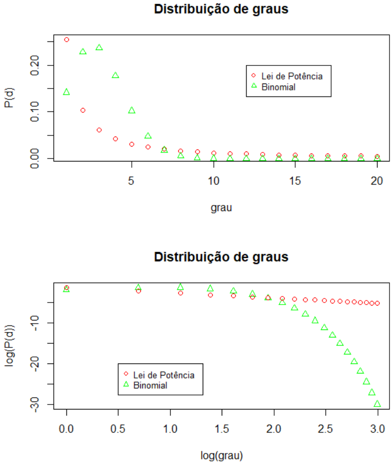


Figura: Distribuição dos Graus em escalas linear e logarítmica, respectivamente, de uma rede aleatória ($n = 31$ e $p = 0{,}1$) e de uma rede lei de potência ($\alpha = 1{,}3$).

Métricas Globais

```
g = nx.karate_club_graph()
a = g.degree(0)                  #grau do nó 0
b = list(g.degree([0,1,2]))  #lista de graus dos nós 0, 1 e 2
c = nx.degree_histogram(g)   #lista com as frequências de ocorrência
                                 #grau 0, grau 1, etc.
print(a)
print(b)
print(c)
```

```
#Resultados
16
[(0, 16), (1, 9), (2, 10)]
[0, 1, 11, 6, 6, 3, 2, 0, 0, 1, 1, 0, 1, 0, 0, 0, 1, 1]
#11 nós têm grau 2
```

Métricas Globais

Coesão: agrupamento

O coeficiente de agrupamento total é dado pela proporção de vezes que dois vértices j e k que são vizinhos de um mesmo vértice i também são vizinhos entre si:

$$cl(G) = \frac{\sum_{(i,j,k):j\neq i,k\neq i,k\neq j} A(i,j)A(i,k)A(j,k)}{\sum_{(i,j,k):j\neq i,k\neq i,k\neq j} A(i,j)A(i,k)}.$$

Métricas Globais

Coesão: agrupamento

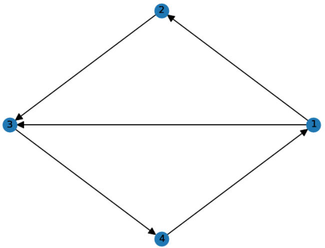
Em redes direcionadas, existem algumas variações de como medir o coeficiente de agrupamento total. A primeira delas, $cl^{d,1}(G)$, é feita ignorando a direção das arestas e calculando este coeficiente como se a rede fosse não direcionada. Um segunda forma seria vendo a proporção de triplas transitivas que ocorrem. Ou seja, das vezes em que o vértice i está conectado ao j e o j está conectado ao k, quantas vezes o i está ligado ao k diretamente:

$$cl^{d,2}(G) = \frac{\sum_{(i,j,k):j\neq i,k\neq i,k\neq j} A(i,j)A(i,k)A(j,k)}{\sum_{(i,j,k):j\neq i,k\neq i,k\neq j} A(i,j)A(j,k)}.$$

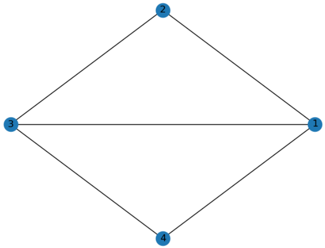
Exemplo

Para calcular o valor de $cl^{d,1}(G)$ da Figura (a), deve-se ignorar a direção das arestas e obter o valor do coeficiente de agrupamento total considerando a Rede Não-Direcionada ilustrada na Figura (b). Deste modo, tem-se que

$$cl^{d,1}(G) = \frac{2+1+2+1}{3+1+3+1} = 0.75.$$



(a) Rede Direcionada.

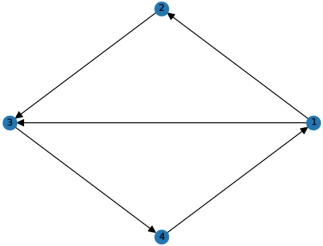


(b) Rede Não-Direcionada Obtida a Partir da Rede Direcionada Ignorando as Direções das Arestas.

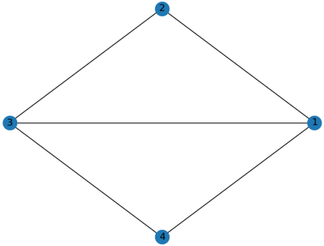
Exemplo

Por outro lado, considerando a direção das arestas o valor de $cl^{d,2}(G)$ é:

$$cl^{d,2}(G) = \frac{1+0+0+0}{2+1+1+1} = 0.2.$$



(a) Rede Direcionada.



(b) Rede Não-Direcionada Obtida a Partir da Rede Direcionada Ignorando as Direções das Arestas.

Métricas Globais

```
g = nx.karate_club_graph()
#círculo com 6 nós
gci = nx.circulant_graph(6,[1])
#grafo regular com 6 nós e 3 arestas por nó
gre = nx.random_regular_graph(3, 6)
#grafo completo com 6 nós
gco = nx.complete_graph(6)

print(nx.transitivity(g))
print(nx.transitivity(gci))
print(nx.transitivity(gre))
print(nx.transitivity(gco))
```

```
#Resultados
0.2556818181818182
0
0.3333333333333333
1.0
```

Métricas Globais

Reciprocidade

Em redes direcionadas, uma medida de interesse é saber qual a fração de arestas que ocorrem em ambas as direções.

$$rc(G) = \frac{\sum_{i,j} A(i,j)A(j,i)}{\sum_{i,j} A(i,j)}.$$

Métricas Globais

```
#Grafo direcionado
gdi = nx.DiGraph()
#Lista de nós
gdi.add_nodes_from(['A','B','C'])
#Arestas
gdi.add_edge('A','B')  #A -> B
gdi.add_edge('B','A')  #B -> A
gdi.add_edge('B','C')  #B -> C
gdi.add_edge('C','A')  #C -> A

re = nx.reciprocity(gdi)  #reciprocidade
print(re)
```

```
#Resultado
0.5
```

Redes de Mundo Pequeno

Ideia

- Observam-se muitas redes complexas grandes que possuem diâmetro e tamanho médio dos caminhos pequenos, o que veio a ser conhecido como redes de mundo pequeno.
- Supondo que uma pessoa adulta tenha cerca de 100 outras pessoas com as quais ela mantém um contato regular. Então, fazendo uma suposição grosseira de que a rede social de conhecidos seja uma árvore, essa pessoa teria 100^2 conhecidos dos seus conhecidos e 100^3 conhecidos dos conhecidos dos seus conhecidos. Ou seja, em 5 passos essa rede atingiria toda população humana na terra.

Redes de Mundo Pequeno

Experimento de Milgram (1967)

Milgram selecionou pessoas desconhecidas de duas cidades americanas e mandou que certas pessoas mandassem cartas para seus conhecidos próximos, pedindo para que eles repassassem as cartas ou para os destinatários alvo ou para conhecidos que pudessem conhecer os destinatários alvo. Eles apenas sabiam o nome e a região aproximada da localização da residência do destinatário. Como resultado do experimento, aproximadamente 1/4 das cartas chegaram ao seu destinatário alvo.